MEASURING THE EFFECT OF TRAINING INTERVENTIONS IS OFTEN DEMANDED BY ORGANISATIONS AND COMMONLY ACCEPTED BY TRAINERS. HOWEVER, THIS IS NOT THE CASE WITH COACHING. THE RESULT OF INDIVIDUAL COACHING IS OFTEN REGARDED AS A PRIVATE, INTERNAL PROCESS. ALSO IT IS INFLUENCED BY MANY UNCON-TROLLABLE OUTSIDE FACTORS, THAT MAKE IT DIFFICULT TO ISOLATE WHAT CHANGE OCCURRED THROUGH COACHING AND WHAT HAPPENED DUE TO OTHER INFLUENCES. THE QUESTION IS ALSO: WHEN CAN COACHING (AS A LONG TERM INTERVENTION) BE REGARDED SUCCESSFUL? IF ONE WANTED TO MEASURE THE EFFECT OF A COACHING INTERVENTION, HOW WOULD THE DESIGN CRITERIA FOR DEVELOPING A MEASUREMENT TOOL LOOK LIKE?

# Measuring Coaching: Selling Hot Air or Breaking a Taboo?

#### BY HEIKE WABBELS AND IDRIES KAHAAR

Attempting to measure the impact of coaching may seem like a betrayal for many coaches - 180 degrees the opposite of what they learned in their coach education, where letting go of the result and focusing on facilitating the process is the learning goal. In corporations that are used to 'measure and control', however, the question is being posed regularly. A private study of the Corporate Leadership Council in 2003 on executive coaching in multinationals (only issued to member organisations) identified 'inconsistent returns' of coaching as one of the key factors that prevent business from incorporating coaching in their leadership development programs. At the same time, professional organisations are trying to improve quality and credibility of the profession through accreditation and standards (EMCC, ICF). In practice, on both sides, 'there is still confusion about standards and terminology, and little evaluation is taking place.'

(www.hrmguide.co.uk/hrd/coaching-survey.htm) Is then the call for evaluation the right solution to turn coaching into a better accredited means of HRD? Or does the very question show the fear of an organisation of not being able to control the personal development of an employee? In a more friendly fashion, such a tool could be a means to increase communication among the stakeholders: organisation, coachee, and coach. But first of all: what would one want to measure in such an evaluation... and how can such a tool be designed?

In 2004, the authors pursued a full-time MBA programme at the University of Cape Town, a Business School that proposes that it is not merely financial know-how and economic insight, but rather the relationships employees have with their organisation and among each other that form the new 'cutting edge' in business. In addition to the traditional MBA courses, classes in 'group facilitation', 'advanced leadership', and 'Integral Coaching' are being offered. The University of Cape Town (UCT) has its own Centre for Coaching that provides corporate coaching and offers courses in the UCT curriculum, based on Integral Coaching theory (see box for information on Integral Coaching). Partly inspired to learn more about coaching, partly adventurous, partly intrigued by the number-crunching courses of the first semester, we decided to attempt to draft a process to measure the effect of coaching as part of our MBA thesis. Our ambitious goal was to prove that it is possible to measure the change that occurrs in a coachee through coaching.

The following article is an account of the research process that resulted in a tool that visualises the effect of a coaching intervention on the coachee. We will introduce 'Integral Coaching', a coaching method taught at UCT, along with 'Ontology', the theory that provided the framework on which our tool was built, and outline the process we followed that resulted in the Integral Coaching Assessment Tool (ICAT). This article is meant to generate further discussion on the subject of measuring the impact of coaching and the process of how measuring could be attempted.

## TURNING JELL-O INTO BUILDING BLOCKS

How does one quantify the impact of something as qualitative as coaching? Isn't that measuring water with a sieve? Every coach and coachee can tell you that effects of a coaching intervention often only show in the long-term when, after a time of reflection and action learning, the new mindset has actually produced new behaviour. Also, the change due to the coaching intervention cannot be isolated from other influences such as organisational culture, economic and political influences or personal experiences. Another challenge is to define objective 'results' of such an intangible experience, a very private and sometimes even unconscious process that may or may not be sustainable. Also, to our knowledge, no academically valid empirical research on measuring the effect of coaching existed. Consequently, we started at the beginning.

In the early days when coaching was introduced in the realm of business, it was discussed in literature as a tool to enhance performance: 'The objectives of coaching are focused on improving individual performance and personal satisfaction, and, consequently, enhancing organizational effectiveness' (Kilburg, 1996, cited in Day, 2001, p. 590). In this context it makes sense that the logical consequence would be to find a correlation among financial return on investment (ROI) in a company that offers personal coaching to their employees. In his article 'Maximizing the Impact of Executive Coaching' Mc Govern consequently proposes that 'Coaching translates into doing. Doing translates into impacting the business. This impact can be quantified and maximized.' (McGovern et al., 2001, p. 1). However, our aim was not to test the financial implications, but rather the internal change that occurs in an individual involved in a coaching process. Or in other words: turn Jell-O into building-blocks. The mould for this had to be:

- · Quantifiable;
- Grounded in theory;
- Provide a benchmark and therefore include a longitudinal aspect (be applicable over time).

And, as the study should be relevant for many organisations and coaches, it should focus on describing a process that could be replicated at different organisations, rather than a 'one size fits all' solution.

# IN WHAT FIELD ARE YOU PLAYING WHEN YOU TALK 'COACHING'?

Through a literature study, we positioned the type of coaching we wanted to research in relation to psychology and behaviour modification and found a matrix that guided our thinking (see figure 1): as coaching is future-oriented, the change will not be linear, but iterative. A tool would be needed that could give insight into the zigzag development of the coachee over time.

Wikipedia.org (http://en.wikipedia.org/wiki/Coaching, January 25, 2007) shows that coaching is being used as an umbrella term, specified by adjectives describing the method or the goal, including (not an exhaustive list): life coaching, career coaching, systemic coaching, ontological coaching, and dissertation coaching. It has been said that there are as many coaching methodologies as there are coaches since each indi-

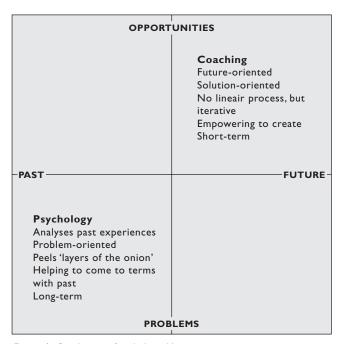


Figure 1: Coaching vs. Psychology Matrix

vidual coach's repertoire is in the end unique. In our particular case we decided to focus on 'Integral Coaching' that was being taught by the School for Coaching at UCT at the time of our MBA studies. Integral Coaching has its theoretical roots in different continents, disciplines and traditions of thought. The American psychologist-philosopher Ken Wilber is the foremost thinker of integral philosophy. Wilber has developed the so-called 'All Quadrants Model' that Flaherty later adapted for Integral Coaching purposes. This model has two axis: 'inside and outside' and 'individual and collective'. Two central theorists in this field are the Chilean academics Humberto Maturana and his student Fernando Flores (Maturana and Varela, 1987). In practice, Integral Coaching has been shaped by James Flaherty and his coaching organisation New Ventures West in the United States.

If 'coaching translates into doing' as McGovern proposes (McGovern et al., 2001), one can consequently presume that 'theory translates into coaching', meaning that every coaching method must have a theoretical underpinning. We therefore researched the theoretical roots of integral coaching further, as

we wanted to craft a tool that was derived from the same theory as the coaching method, and found another source of influence on Integral Coaching: Ontology (meaning 'way of being'), that has emerged from the integration of significant 20th-century developments in the biology of cognition, existential philosophy, and the philosophy of language. Fernando Flores, Rafael Echeverria, and Julio Olalla (Ollala, 2004) have played an important role of crafting the theoretical framework for ontological coaching, so has the Australian Alan Sieler (Sieler, 2003). Sieler points out that behaviour 'is determined by how we are observing situations. The particular perspective or "angle" we have of a situation, determines how we will respond and deal with what is occurring (Sieler, 2003, p. 33).' This belief, that the 'structure of interpretation' of the coachee shapes how they act in the world, was the turning key to our research process: while we could not measure what exactly causes a change, we can very well assess how the coachee feels about themselves before and after the coaching - and compare this to actual observable behaviour.

Since Integral Coaching works on the level of the beliefs and perception of the coachee ('structure of interpretation') that is the cause for changes in behaviour, the challenge lays in the question of how to assess a subjective item like a changed belief, an increased feeling of well-being, a stronger motivation in mutually understood language. As the goal of Integral Coaching is "long-term excellent performance, self-correction, and self-generation" (Flaherty, 1999), it seemed logical to describe the very drivers of behaviour, the beliefs, and craft them into statements. We felt we could capture this best in a questionnaire.

# THE PROCESS: BRINGING THE INTEGRAL COACHING ASSESSMENT TOOL (ICAT) TO LIFE

In the further process we adopted a positivist approach, meaning that we based our research on measurable, quantifiable observations. We included both inductive and deductive approaches as we moved from theory to data into designing and testing the questionnaire

In designing a questionnaire to test our hypothesis we followed the development recommendations as outlined by Churchill and Saunders (Saunders, 2003), see figure 3.

# **Textbox I: Ontology**

Sieler argues that human behaviour is determined by how a person perceives the world and that this perception is reflected in the individual's use of language, their physiology and their emotional experiences. According to him, sustainable behavioural change can be stimulated through a coach who can assist the coachee with observing how he/she unconsciously perceives the world. In the related field of Integral Coaching, Flaherty refers to an individual's perception as their 'structure of interpretation' (Flaherty, 1999, p. 8). Thus, changes in this structure will eventually lead to changes in behaviour.

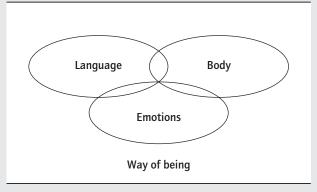


Figure 2: (Adapted from Sieler, 2003, p. 34)

### Step 1: Specify domain of construct

First, we assessed how our hypothesis could be tested. We came to the conclusion that the only objectively measurable factor in coaching is the behaviour of the coachee. This might seem contradictory, since Integral Coaching is not a behaviour modification. We therefore do not imply that certain behaviour should change, but we see it as an indicator that the coachee has undergone an internal process that caused them to adapt their behaviour. In the context of this research, behaviour is therefore understood as the ocular result of the individual's internal beliefs. We continued the process by defining which behaviour we wanted to include. We used the structure that integral theory provided, the four quadrants:

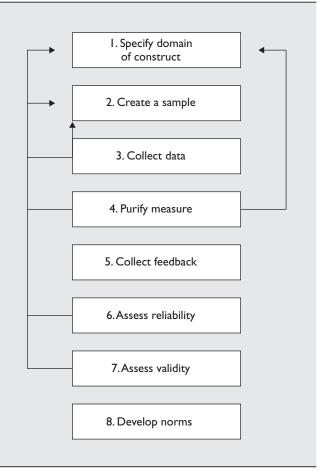


Figure 3: Questionnaire development stages

- Quadrant I represents all aspects of the nervous system, like moods and sensory input, of a person within themselves: it assesses the emotional level: 'inside me'. Examples of domains: thoughts, feelings, moods
- Quadrant II focuses on the interaction of a person in terms of their body and behaviour 'outside me'. Exam-ples of domains: neuro-muscular system, somatics, behaviour.
- Quadrant III represents the relationships of a person as a member of a group: 'inside group'. Examples of domains: culture, rituals, language, history, customs.
- Quadrant VI deals with the interaction of a person with their environment: 'outside group'. Examples of domains: physical laws, technology, process, opjects.

In the next step, we defined qualities per quadrant building on previously developed definitions by O'Flaherty and Flaherty (no family, Flaherty, 1999). We developed this further in a series of workshops with two coaches from the Centre for Coaching of UCT who were experienced with the method and the literature in the field of integral theory and ontology. The guiding question was "What qualities are typical and represent the concept of the quadrant?" This exercise resulted in 20 qualities we wanted to assess in the questionnaire (see figure 4).

Now, the challenge resided in making the step from an intangible concept to a common language that allows the subjective attitudes to become 'inter-subjective'. Practically: what exactly is meant by 'empathy', or 'rigour', or 'confidence'? This resulted in a 'Qualities and Behaviour Matrix' (see figure 8) that we see as the core process-step of our research. We compiled interpretations from literature on the respective qualities. For this, collected quotes literature on integral theory, ontology, philosophy, and psychology. The reason for this was that some qualities, like 'balance', 'trust' or

Theory translates into coaching methods and coaching translates into doing.

'objectivity' have many different facets, and it should be exactly clear what aspect of a quality was being meant. Following this, the theory was translated into behavioural and attitudinal statements. The guiding question was: 'What behaviour and belief would a person who has mastered this quality typically show?'

As there are four quadrants, we wanted to create four statements per quality: one from the of each quadrant. So we would test the quality in all four domains: regarding the thoughts and feelings of individual, regarding the body and

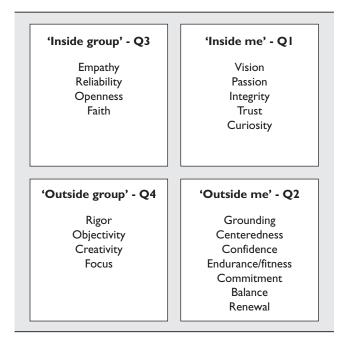


Figure 4: Defining qualities per quadrant

behaviour, regarding the environment, and regarding their role in a group. This accumulated to a total of 80 questions. We regarded this as the compromise between thoroughness and the goal to keep the questionnaire manageable for the respondent. In some cases extra questions were created for testing purposes.

# Step 2: Create a sample

In the second step we identified a sample. Saunders suggests to test a questionnaire for a student research with a minimum of ten individuals (Saunders et. al., 2003, p. 309). The questionnaire was tested with fellow MBA students, as well as managers at an international organisation in the tobacco industry, where Integral Coaching was part of their leadership development programme.

# Step 3: Collect data / 4: Purify measure

In step 3 we consequently collected and computerised the responses to the questionnaire as well as the comments on the feedback sheet we provided along with the test-questionnaire. On the basis of this input we revised the questionnai-

re (step 4). For example, we took out ambiguous questions and adapted the language, as well as the lay-out where this was not clear.

Step 5, 6, 7: Collecting data, assessing reliability and validity While we did perform a pilot test and collected feedback (step 5), we did not test for reliability and validity (steps 6, 7), as this needs a bigger sample. The results of the pilot test, however, gave great insights in the practicality. We tested with 26 persons: two coaches, 16 executives at a South African multinational who had undergone Integral Coaching in as part of a leadership development programme, and ten MBA students as 'control group'. The pilot resulted in valuable feedback. We, for instance, extended the number of grade points from four to ten to create a finer rating grid to suggest that there is more room for development. It also became clear that, in its current form, the framework is not suitable for comparison among coachees, but is a strictly personal assessment tool, much like the Enneagram. Also, further research with a bigger sample is needed to develop norms (step 8). With a sample

The particular perspective or angle we have of a situation, determines how we respond and deal with what is occurring.

of n > 100 (Saunders et al. 2003, p. 309) norms could be indicated, such as typical phases in development, averages, and correlations within the same age group.

# CRUNCHING NUMBERS: HOW DOES ONE VISUALISE A CHANGED BELIEF?

In the same phase we designed a Questionnaire Rating System (QRS) that we wanted to test together with the first draft of the questionnaire. The QRS is an Excel application that weighs the answers to the questionnaire, so that the results of the respective quadrants are in balance, as some quadrants

have more qualities than others (4-4-5-7). After a respondent has completed the questionnaire the coach can enter the respondent's scores into the QRS Excel sheet. The respondent's weighted scores per quadrant will then be automatically calculated within the spreadsheet. These scores can then be plotted in a graph that illustrates the score and discussed with the coachee (see figure 5).

#### **CONCLUSION AND APPLICATION**

In this article we have outlined the research process that resulted in a tool to visualise the effect of coaching on a coachee. The challenge was to define parameters to measure the result of a coaching intervention. This exercise should be quantifiable, grounded in theory, provide a benchmark, and include a longitudinal aspect. The article is an account of the process. As a central part of the research, we funnelled the theory of Integral Coaching and Ontology to describe observable behaviour. The result was a questionnaire with 80 questions, the Integral Coaching Assessment Tool (ICAT). To visualise and weigh the answers, we developed the questionnaire rating system (QRS). In the process of

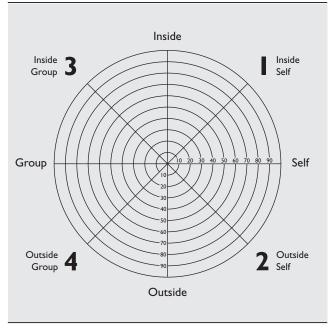


Figure 5: Questionnaire Rating System plotted in a graph

# **Textbox 2: Qualities and Behaviour Matrix**

Even when two people apply the same quality, they will have their unique way of demonstrating it. Two different individuals can have two different ways of being patient: one person can be patient with people in a group, while the other is patient with themselves, but might not necessarily be patient with things around them (e.g. being patient with starting the car, or letting people finish their sentences, versus being patient with one's own personal development).

So, in assessing beliefs one enters a completely subjective realm; in order to provide some standard, we needed to create a com-

mon language that made it possible to describe the subjectively observed belief in a common language. We started the process of crafting the 'Qualities and Behaviours Matrix' by compiling definitions, and fragments of authors in the field of integral theory. In order to create a tool that was true to its theoretical roots, we felt we had to test every aspect of a quality, since a quality can be 'perceived' (quadrant 1), 'embodied' (quadrant 2), 'engaged' (quadrant 3), and 'enacted' (quadrant 4).

We therefore crafted behaviours statements from the angle of each of the four quadrants:

Nr.	Q	Quality - Def.	Context of Quality	Quadrant I (percieve)	Quadrant 2 (embody)	Quadrant 3 (engage)	Quadrant 4 (enact)
14	3	the state of being consistent.	'When we apply the same standards to ourselves that we apply to our clients we are being self-consistent.' (Flaherty 1999, p. 152). 'Being able to rely on people is indispensable for developing constructive and productive relationships. But what does it mean when we say that we can rely on someone - what exactly do we find reliable? Two key dimensions of reliability are time and standards. Continually arriving late for meetings, not meeting agreed deadlines and keeping others waiting on social occasions are examples of behaviour that generates mistrust. Similarly with standards.' (Siegler, 2003, p. 342).	(reliability and dependability).	yl am reliable and dependable.	I apply the same standards to myself as I do to others.	I structure my environment in a way that a allows me to be reliable and dependable (e.g. I use an organiser).

Figure 6: Qualities and Behaviours Matrix

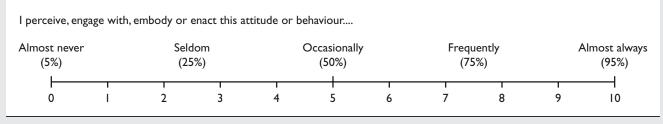


Figure 7: Questionnaire scale

#### CONSISTENCY

- 1. I value consistency (reliability and dependability).
- 2. I am reliable and dependable.
- 3. I apply the same standards to myself as I do to others.
- I structure my environment in a way that allows me to be reliable and dependable (e.g. I use an organizer).

### INTEGRITY

- I feel that it is important to live a life driven by values.
- 2. My values show in my actions and behaviour.
- 3. I do not compromise my values for others.
- 4. I have a system in place in order to uphold my values (e.g. going to church).

Figure 8: Examples of questions on the qualities Consistency and Integrity per quadrant. In the actual questionnaire, they are scrambled.

our research we learned to mould the theoretical underpinning behind a practical coaching method into the academic steps necessary to draft a questionnaire. In the end, our conclusion is that it is possible to measure the effect of coaching.

In the course of our research, it became clear that although such a questionnaire can be a very valuable starting point for coaches and coachees (see the paragraph on 'application' below), for us it was rather the process of defining the behaviour statements that was valuable. We gained a much deeper understanding of the theory and sharpened our eye to observe behaviour. The question is: how could this experience be offered to others?

It is not important to achieve a 'high' score or a 'low' score, but the score is an individual snapshot in time.

### SOME POSSIBLE AREAS OF APPLICATION

Coach and Coachee. As with the Enneagram, the purpose of an evaluation like this is tracking the coachee's personal development; it is not a tool to compare one coachee with another (yet! – with a bigger sample this could be an interesting future step). What we personally like about it is that it is not important to achieve a 'high' score or a 'low' score, but that the score is an individual snapshot in time that differs from person to person. The questionnaire assessment could serve as a basis for a first meeting between the coachee and coach. At the end of a coaching engagement, it can be used to debrief and assess the change that has occurred. Also, it gives the coachee a visual image of the change in their attitude towards their own behaviour.

The Organisation. This tool can be used by anyone using integral coaching. However, we think that it is actually the definition phase (crafting the statements for the question-

naire, e.g. What is integrity? What is consistency?) that would be of most benefit to an organisation. Defining the statements forced us to have a good discussion of what we actually see and think when we observe a behaviour. If this conversation was held in a management team, this process would help create a mutual language, more ownership of the process and more focus on desired behaviour in the organisation.

South African companies are still very sceptical of coaching and have often never heard of the concept.

After all, the goal should not be to have another tool in the toolbox, but to improve relationships and co-creation at work.

**Coach Training.** Crafting a matrix with definitions and qualities can be used in coach training to give the coachee a deeper understanding of the concepts meant by a specific coaching method.

The Coaching Profession. As the field has evolved through practice, more academic discussion and research is needed. We hope that this article inspires fellow coaches and academics to continue the discussion on 'the measurement of coaching' and take this as a starting point to refine or develop own models.

# REFLECTION ON DUTCH AND SOUTH AFRICAN EXPERIENCE WITH MEASURING COACHING

Over the past two years, we have gained practical experience in coaching on two continents: in Africa and Europe. We experience very different attitudes towards measuring coaching. That is why both of us reflect on our own experience with 'measuring coaching' from a South African and a Dutch perspective:

Idries' Conclusion: A South African Perspective

# Measuring The Business Impact Is In Demand

South Africa is a fledgling democracy with a broad range of social, political and economic challenges. These include but are not limited to HIV/AIDS, poverty and mass unemployment (50%+) mainly the result of 50 years of Apartheid. Within the business arena there is a focus on Black Economic Empowerment and gender equality via initiatives such as recently introduced employment equity laws and affirmative action policies. This makes 'harnessing diversity' particularly important to South African companies and their ongoing survival is becoming more and more dependant on it.

Here awaits an opportunity for coaches to make a significant difference. Coaches that are able to demonstrate this difference in a tangible way (through effective measurement) will be in high demand in South Africa.

Local companies are still very sceptical of coaching and have often never heard of the concept. This makes it very challenging to sell coaching services locally and this, in turn, makes it especially important that coaches are able to offer prospective clients decent proof of past successes, hence the importance of measurement. However, underlying every business coaching goal is at least one personal development need. At first glance it may therefore appear that there is more of a need to measure the business value of coaching than there is a need to measure the individual's change within a personal development context.

Effective measurement within a personal development context is an important aspect of organizational coaching in South Africa. However, it is important for the coach to clearly describe the link between business impact and personal development to prospective clients so that clients understand why it is important to focus on and thus measure the change in an individual from a personal development perspective. In the end it is the results that speak loudest and this directly impacts on a coach's credibility.

#### Heike's Conclusion: A Dutch Perspective

### Getting People in Motion, Don't Perfect Them

The Dutch are quite experienced with executive and personal coaching. That might be one reason why the demand for evaluation and return on investment has not really been an issue in my practice. Instead, much care is given to the selection of a good coach. In 2007, coaching is a well-accepted means to support personal development on and off the job. Many organisations have realised that in order to be successful, professionals need hard- and soft skills and that soft skills need time to develop. Progress, not perfection is the goal.

Also, diversity is an issue that becomes not only more pressing in Holland, but with the continuing expansion of Europe for the entire continent. Organisations realise that in order to stay innovative and flexible, they need to foster diversity and share knowledge. This requires 'old-school' managers to re-examine their beliefs about potential, talent and collaboration. It also made coaching in the business setting more acceptable.

As the job market is 'tight' (meaning there is not enough qualified personnel), coaching is frequently used in HRD as a means to attract and bind employees to the organisation. Or, the opposite, use coaching to increase internal mobility. Altogether, the concept of life-long-learning has entered the realm of business. Coaching is therefore seen as a means to 'get people in motion' and to support a learning organisation.

Concluding, I find that we did not break a taboo, nor inflate a hot-air balloon, I rather gained the insight that it is neither the tool nor the result of the measuring, but rather the way coaching is embedded in an organisation that is most important for the success of it. That would mean turning the mirror the other way: not to measure the results, but to assess why the question for measuring is put forward in the first place: is the wish to measure rather a sign of lack of trust? Lack of ownership? The wish to control?

**Idries Kahaar** works as an executive coach in Cape Town, South Africa. His work includes coaching within a leadership development as well as entrepreneurship context where he assists business owners and managers with learning their way out of strategic problems.

Heike Wabbels is an associate with Kessels & Smit, The Learning Company, in the Netherlands. She advises organisations on questions regarding learning and talent development and works with individuals and groups on potential development and career planning.

### **REFERENCES**

Day, D.V., (2001). Leadership Development: A Review in Context. *Leadership Quarterly*, Vol 11, pp. 581 - 613.

Flaherty, J. (1999). *Coaching: Evoking excellence in others*. Boston: Butterworth Heinemann.

Kahaar, I. and Wabbels, H. (2004). If, and under what conditions, does the use of coaching in leadership development add value to a leadership development program of an organization using British American Tobacco in South Africa as an example. Development of a framework for measuring the value-add of coaching in leadership development. Cape Town: GSB UCT.

McGovern, J., Lindemann, M., Vergara, M., Murphy, S., Barker, L. and Warrenfeltz, R. (2001). Maximizing the Impact of Executive Coaching. *Manchester Review*, Vol 6, pp. 1-9.

Maturana, H. and Varela, F.. (1987). *The Tree of Knowledge*. Boston: Shambhala Publications.

Ollala, J. (2004). From Knowledge to Wisdom. Boulder, CO: Newfield Network.

Sieler, A. (2003). *Coaching to the Human Soul.* Victoria: Newfield Australia.

Saunders, M., Lewis, P. and Thornhill, A. (2003). Research Methods for Business Students. Essex: Pearson Education Limited.

#### WERSITES

www.worldofkenwilber.com
www.hrmguide.co.uk/hrd/coaching-survey.htm
www.newventureswest.com
http://en.wikipedia.org/wiki/Coaching
(All websites were accessible at the time of print of this article.)